

OPTIMIZED ESTIMATES OF THE REGULARITY OF THE CONDITIONAL DISTRIBUTION OF THE SAMPLE MEAN

VICTOR CHULAEVSKY

ABSTRACT. We give an improved estimate for the regularity of the conditional distribution of the empiric mean of a finite sample of IID random variables, conditional on the sample "fluctuations", extending the well-known property of Gaussian IID samples. Specifically, we replace the bounds in probability, established in our earlier works, by those in distribution, and this results in the optimal regularity exponent in the final estimate.

1. INTRODUCTION

Consider a sample of N IID (independent and identically distributed) random variables with Gaussian distribution $\mathcal{N}(0, 1)$, and introduce the sample mean $\xi = \xi_N$ and the "fluctuations" η_i around the mean:

$$\xi_N = \frac{1}{N} \sum_{i=1}^N X_i, \quad \eta_i = X_i - \xi_N, \quad i = 1, \dots, N.$$

It is well-known from elementary courses of the probability theory that ξ_N is independent from the sigma-algebra \mathfrak{F}_η generated by $\{\eta_1, \dots, \eta_n\}$ (the latter are linearly dependent, and have rank $N - 1$). To see this, it suffices to note that η_i are all orthogonal to ξ_N with respect to the standard scalar product in the linear space formed by X_1, \dots, X_N given by

$$\langle Y, Z \rangle := \mathbb{E}[Y Z],$$

where Y and Z are real linear combinations of X_1, \dots, X_N (recall: $\mathbb{E}[X_i] = 0$).

Therefore, the conditional probability distribution of ξ_N given \mathfrak{F}_η coincides with the unconditional one, so $\xi_N \sim \mathcal{N}(0, N^{-1})$, thus ξ_N has bounded density

$$p_\xi(t) = \frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi N^{-1}}} \leq \frac{N^{1/2}}{\sqrt{2\pi}}.$$

Moreover, for any interval $I \subset \mathbb{R}$ of length $|I|$, we have

$$\text{ess sup } \mathbb{P} \{ \xi_N(\omega) \in I \mid \mathfrak{F} \} = \mathbb{P} \{ \xi_N(\omega) \in I \} \leq \frac{N^{1/2}}{\sqrt{2\pi}} |I|. \quad (1.1)$$

The essential supremum in the above LHS is a bureaucratic tribute to the formal rule saying that $\mathbb{P} \{ \cdot \mid \mathfrak{F} \}$ is a random variable (which is \mathfrak{F} -measurable), and as such is defined, generally speaking, only up to subsets of measure zero.

In some applications to the eigenvalue concentration estimates in the theory of multi-particle random, Anderson-type Hamiltonians, one has to estimate the

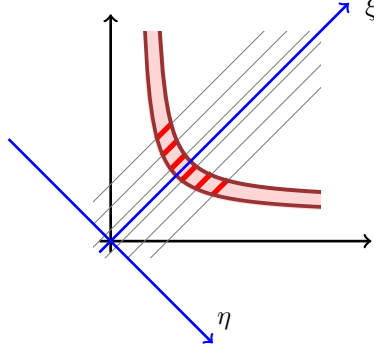


FIGURE 1. In this example, $N = 2$, $\xi = \frac{1}{2}(X_1 + X_2)$ and $\eta = \frac{1}{2}(X_1 - X_2)$. One has to assess the probability of the pink curvilinear strip $\{(X_1, X_2) : \xi \in [a(\eta), a(\eta) + s]\}$.

probability of the form

$$\mathbb{P}\{\xi_N(\omega) \in I(\eta)\},$$

where the interval $I(\eta) = [f(\eta), f(\eta) + \epsilon]$ is determined only by the fluctuations η_\bullet , and f is some measurable (in fact, Lipschitz continuous¹) function. For example, with $N = 2$,

$$\xi = \xi_2 = \frac{X_1 + X_2}{2}, \quad \eta = \eta_1 = \frac{X_1 - X_2}{2},$$

one may consider the probability

$$\mathbb{P}\{\xi \in [\eta^2, \eta^2 + s]\} = (2\pi)^{-1} \int_{\mathbb{R}^2} dX_1 dX_2 e^{-\frac{1}{2}(x_1^2 + x_2^2)} \mathbf{1}_A(x_1, x_2)$$

where, e.g.,

$$A := \left\{ (x_1, x_2) \in \mathbb{R}^2 : \frac{(x_1 - x_2)^2}{4} \leq \frac{x_1 + x_2}{2} \leq \frac{(x_1 - x_2)^2}{4} + s \right\}, \quad s > 0.$$

In this particular case – for Gaussian samples – the conditional regularity of the sample mean ξ_N (given the fluctuations) is granted, but is not always so, as shows the following elementary example where the common probability distribution of the sample X_1, X_2 is just excellent: $X_i \sim \text{Unif}([0, 1])$, so X_i admit a compactly supported probability density bounded by 1. In this simple example the random vector (X_1, X_2) is uniformly distributed in the unit square $[0, 1]^2$, and the condition $\eta = c$ selects a straight line in the two-dimensional plane with coordinates (X_1, X_2) , parallel to the main diagonal $\{X_1 = X_2\}$. The conditional distribution of ξ given $\{\eta = c\}$ is the uniform distribution on the segment

$$J_c := \{(x_1, x_2) : x_1 - x_2 = 2c, 0 \leq x_1, x_2 \leq 1\}$$

of length vanishing at $2c = \pm 1$. For $|2c| = 1$, the conditional distribution of ξ on J_c is concentrated on a single point, which is the ultimate form of singularity.

¹We refer to the applications where f is an eigenvalue of some self-adjoint operator, and by the min-max principle, such EVs are Lipschitz continuous functions of the parameters upon which the operator depends.

2. AN APPLICATION TO THE WEGNER-TYPE BOUNDS

Let Λ be a finite graph, with $|\Lambda| = N \geq 1$, and $H(\omega) = H_\Lambda(\omega)$ be a random DSO acting in the finite-dimensional Hilbert space $\mathcal{H} = \mathcal{H}_\Lambda = \ell^2(\Lambda)$, with IID random potential $V : \Lambda \times \Omega \rightarrow \mathbb{R}$, relative to a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. Decomposing the random field V on Λ ,

$$V(x; \omega) = \xi_N(\omega) + \eta_x(\omega),$$

we can represent $H(\omega)$ as follows:

$$H(\omega) = \xi_N(\omega) \mathbf{1} + A(\omega),$$

where the self-adjoint operator $A(\omega)$ is \mathfrak{F}_η -measurable, and so are its eigenvalues $\tilde{\mu}_j(\omega)$, $j = 1, \dots, N$. It is readily seen that $A(\omega)$ is a DSO with potential having zero sample mean. Since $A(\omega)$ commutes with the scalar operator $\xi_N(\omega) \mathbf{1}$, the eigenvalues $\lambda_j(\omega)$ of $H(\omega)$ have the form

$$\lambda_j(\omega) = \xi_N(\omega) + \mu_j(\omega). \quad (2.1)$$

The numeration of the eigenvalues $\lambda_j(\omega)$, $\mu_j(\omega)$ is, of course, not canonical, but they can be consistently defined as random variables on Ω .

The representation (2.1) implies immediately the following EVC bound: for any interval $I_s = [t, t + s]$,

$$\begin{aligned} \mathbb{P} \{ \text{tr } P_{I_s}(H(\omega)) \geq 1 \} &\leq \sum_{j=1}^N \mathbb{P} \{ \lambda_j(\omega) \in I_s \} = \sum_{j=1}^N \mathbb{P} \{ \xi_N(\omega) + \mu_j(\omega) \in I_s \} \\ &= \sum_{j=1}^N \mathbb{E} [\mathbb{P} \{ \xi_N(\omega) + \mu_j(\omega) \in I_s \mid \mathfrak{F}_\eta \}] \\ &= \sum_{j=1}^N \mathbb{E} [\mathbb{P} \{ \xi_N(\omega) \in [-\mu_j(\omega) + t, -\mu_j(\omega) + t + s] \mid \mathfrak{F}_\eta \}] \end{aligned} \quad (2.2)$$

Further, omitting the argument ω for notational brevity, we have

$$\begin{aligned} \mathbb{P} \{ \xi_N + \tilde{\mu}_j \in I_s \mid \mathfrak{F}_\eta \} &= \mathbb{P} \{ \xi_N \in [\mu_j + t, \mu_j + t + s] \mid \mathfrak{F}_\eta \} \\ &= \mathbb{P} \{ \xi_N \in [\tilde{\mu}_j, \tilde{\mu}_j + s] \mid \mathfrak{F}_\eta \} \end{aligned}$$

where $\tilde{\mu}_j(\omega) := -\mu_j(\omega) + t$ are \mathfrak{F}_η -measurable, i.e., fixed under the conditioning. Now introduce the conditional continuity modulus of ξ_N , given \mathfrak{F}_η :

$$\nu_N(s) := \sup_{t \in \mathbb{R}} \text{ess sup } \mathbb{P} \{ \xi_N \in [t, t + s] \mid \mathfrak{F}_\eta \}, \quad s > 0.$$

Obviously,

$$\mathbb{P} \{ \lambda_j \in I_s \mid \mathfrak{F}_\eta \} \leq \nu_N(s),$$

thus the unconditional probability $\mathbb{P} \{ \lambda_j \in I_s \}$ can be assessed by analyzing the probability distribution of the random conditional continuity modulus $\nu_N(s; \omega)$.

In this section, we discuss by way of example the Wegner-type bounds for a conventional, single-particle DSO, but in applications to the multi-particle EVC bounds, similar objects turn out to be of interest:

$$s \mapsto \mathbb{P} \{ \xi_N(\omega) \in [\tilde{\mu}(\omega), \tilde{\mu}(\omega) + s] \}, \quad (2.3)$$

where an \mathfrak{F}_η -measurable random variable $\tilde{\mu}$ is given by an eigenvalue of yet another operator $\tilde{H}(\omega)$ which is not necessarily independent of $H(\omega)$. The most difficult case is where $H(\omega)$ and $\tilde{H}(\omega)$ are stochastically correlated in a very strong way: every "local" random variable, representing the disorder in a multi-particle Anderson model, which affects $H(\omega)$ also affects $\tilde{H}(\omega)$, and vice versa. As a result, there is little one can say about $\tilde{\mu}(\omega)$, except that it is a measurable function.

3. REDUCTION TO THE LOCAL ANALYSIS IN THE SAMPLE SPACE

Assume that the support $\mathcal{S} \subset \mathbb{R}$ of the common *continuous* marginal probability measure \mathbb{P}_V of the IID random variables X_j , $1 \leq j \leq N$, is covered by a finite or countable union of intervals:

$$\mathcal{S} \subset \cup_{k \in \mathcal{K}} J_k, \quad \mathcal{K} \subset \mathbb{Z}, \quad J_k = [a_k, b_k], \quad a_{k+1} \geq b_k.$$

Let $\mathbf{K} = \mathcal{K}^N$, and for each $\mathbf{k} = (k_1, \dots, k_N) \in \mathbf{K}$, denote

$$\mathbf{J}_{\mathbf{k}} = \times_{i=1}^N J_{k_i}.$$

Owing to the continuity of the marginal measure, J_k are "essentially" disjoint: for all $k \neq l$, $\mathbb{P}_V(J_k \cap J_l) = 0$. Respectively, the family of the parallelepipeds $\{\mathbf{J}_{\mathbf{k}}, \mathbf{k} \in \mathbf{K}\}$ forms a partition \mathcal{K} of the sample space, which we will often identify with the probability space Ω . Further, let $\mathfrak{F}_{\mathcal{K}}$ be the sub-sigma-algebra of \mathfrak{F} generated by the partition \mathcal{K} . Now the quantities of the general form (2.3) can be assessed as follows:

$$\begin{aligned} \mathbb{P}\{\xi_N \in [\tilde{\mu}, \tilde{\mu} + s]\} &= \mathbb{E}\left[\mathbb{P}\{\xi_N \in [\tilde{\mu}, \tilde{\mu} + s] \mid \mathfrak{F}_{\mathcal{K}}\}\right] \\ &= \sum_{\mathbf{k} \in \mathbf{K}} \mathbb{P}\{\mathbf{J}_{\mathbf{k}}\} \mathbb{P}\{\xi_N \in [\tilde{\mu}, \tilde{\mu} + s] \mid \mathbf{J}_{\mathbf{k}}\}. \end{aligned}$$

Let $\mathbb{P}_{\mathbf{k}}\{\cdot\}$ be the conditional probability measure, given $\{X \in \mathbf{J}_{\mathbf{k}}\}$, $\mathbb{E}_{\mathbf{k}}[\cdot]$ the respective expectation, and $p_{\mathbf{k}} = \mathbb{P}\{\mathbf{J}_{\mathbf{k}}\}$. Then we have

$$\begin{aligned} \mathbb{P}\{\xi_N \in [\tilde{\mu}, \tilde{\mu} + s]\} &= \sum_{\mathbf{k} \in \mathbf{K}} p_{\mathbf{k}} \mathbb{E}_{\mathbf{k}}\left[\mathbb{P}_{\mathbf{k}}\{\xi_N \in [\tilde{\mu}, \tilde{\mu} + s] \mid \mathfrak{F}_\eta\}\right] \\ &\leq \sup_{\mathbf{k} \in \mathbf{K}} \mathbb{E}_{\mathbf{k}}\left[\mathbb{P}_{\mathbf{k}}\{\xi_N \in [\tilde{\mu}, \tilde{\mu} + s] \mid \mathfrak{F}_\eta\}\right]. \end{aligned} \tag{3.1}$$

This simple formula shows that one may seek a satisfactory upper bound on the LHS of (3.1) by assessing the "local" conditional probabilities $\mathbb{P}_{\mathbf{k}}\{\xi_N \in [\tilde{\mu}, \tilde{\mu} + s] \mid \mathfrak{F}_\eta\}$, where each random variable X_j is restricted to a subinterval J_{k_j} of its global support, so the entire sample $X = (X_1, \dots, X_N)$ is restricted to a parallelepiped $\mathbf{J} \subset \mathbb{R}^N$.

In the next section, we perform such analysis first in the case of a uniform marginal distribution of the IID variables X_i .

4. UNIFORM MARGINAL DISTRIBUTIONS

Let be given a real number $\ell > 0$ and an integer $N \geq 2$. Consider a sample of N IID random variables with uniform distribution $\text{Unif}([0, \ell])$, and introduce again the sample mean $\xi = \xi_N$ and the "fluctuations" η_i around the mean:

$$\xi_N = \frac{1}{N} \sum_{i=1}^N X_i, \quad \eta_i = X_i - \xi_N.$$

For the purposes of orthogonal transformation $(X_1, \dots, X_n) \mapsto (\tilde{\xi}_N, \tilde{\eta}_2, \dots, \tilde{\eta}_N)$, we also need a rescaled empirical mean

$$\tilde{\xi}_N = N^{1/2} \xi_N,$$

so

$$X_i = \eta_i + N^{-1/2} \tilde{\xi}_N, \quad i = 1, \dots, N. \quad (4.1)$$

Further, consider the Euclidean space $\sim \mathbb{R}^N$ of real linear combinations of the random variables X_i with the scalar product $\langle X', X'' \rangle = \mathbb{E}[X' X'']$. Clearly, the variables $\eta_i : \mathbb{R}^N \rightarrow \mathbb{R}$ are invariant under the group of translations

$$(X_1, \dots, X_N) \mapsto (X_1 + t, \dots, X_N + t), \quad t \in \mathbb{R},$$

and so are their differences $\eta_i - \eta_j \equiv X_i - X_j$, $1 \leq i < j \leq N$. Introduce the variables

$$Y_i = \eta_i - \eta_N, \quad 1 \leq i \leq N-1, \quad (4.2)$$

Then the space \mathbb{R}^N is fibered into a union of affine lines of the form

$$\begin{aligned} \tilde{\mathcal{X}}(Y) &:= \{X \in \mathbb{R}^N : \eta_i - \eta_N = Y_i, i \leq N-1\} \\ &:= \{X \in \mathbb{R}^N : X_i - X_N = Y_i, i \leq N-1\}, \end{aligned} \quad (4.3)$$

labeled by the elements $Y = (Y_1, \dots, Y_{N-1})$ of the $(N-1)$ -dimensional real vector space $\mathbb{Y}^{N-1} \cong \mathbb{R}^{N-1}$. Set

$$\mathcal{X}(Y) = \tilde{\mathcal{X}}(Y) \cap \mathbf{C}_1 = \{X \in \mathbf{C}_1 : X_i - X_N = Y_i, i \leq N-1\}$$

and endow each nonempty interval $\mathcal{X}(Y) \subset \mathbb{R}^N$ with the natural structure of a probability space inherited from \mathbb{R}^N :

- if $|\mathcal{X}(Y)| = 0$ (an interval reduced to a single point), then we introduce the trivial sigma-algebra and trivial counting measure;
- if $|\mathcal{X}(Y)| = r > 0$, then we use the inherited structure of an interval of a one-dimensional affine line and the normalized measure with constant density r^{-1} with respect to the inherited Lebesgue measure on $\mathcal{X}(Y)$.

The transformation $X \mapsto (\xi_N, \eta_1, \dots, \eta_{N-1})$ is non-degenerate, but not orthogonal. We will have to work with the metric on $\mathcal{X}(Y)$, induced by the standard Riemannian metric in the ambient space \mathbb{R}^N ; to this end, introduce an orthogonal coordinate transformation in \mathbb{R}^N , $X \mapsto (\tilde{\xi}_N, \tilde{\eta}_1, \dots, \tilde{\eta}_{N-1})$, such that

$$\tilde{\xi}_N = N^{-1/2} \sum_{i=1}^N X_i = N^{1/2} \xi_N; \quad (4.4)$$

the exact form of $\tilde{\eta}_j$, $j = 1, \dots, N-1$ is of no importance, provided that the transformation is orthogonal.

Remark 4.1. For later use, note that, owing to (4.4), each of the re-scaled variables $N^{1/2} X_i$ can serve as the (normalized) length parameter on the elements $\mathcal{X}(Y)$. Along an element $\mathcal{X}(Y)$, one can simultaneously parameterize $\tilde{\xi}$ and the variables X_i , by setting $\tilde{\xi}(t) = c_0 + t$, $X_j(t) = c_j + N^{-1/2} t$, with arbitrarily chosen constants c_j . Here, $\tilde{\xi}_N$ is a natural length parameter on $\mathcal{X}(Y)$, since the transformation $X \mapsto (\tilde{\xi}_N, \tilde{\eta}_1, \dots, \tilde{\eta}_{N-1})$ is orthogonal.

It follows from (4.4) that for any given $a \in \mathbb{R}$, $s > 0$, and some $a' \in \mathbb{R}$,

$$\xi_N \in [a, a + s] \iff \tilde{\xi}_N \in [a', a' + N^{1/2}s] \quad (4.5)$$

Next, denote $\mathbf{J}^{(\ell)} = [0, \ell]^N$ and introduce the random variable

$$\nu_N(s; \mathbf{J}^{(\ell)}) = \nu_N(s; \mathbf{J}^{(\ell)}; X) := \text{ess sup} \sup_{t \in \mathbb{R}} \mathbb{P} \{ \xi_N \in [t, t + s] \mid \mathfrak{F}_\eta \}. \quad (4.6)$$

Here the presence of ess sup is the tribute to the fact that the conditional probabilities are random variables, usually defined up to subsets of zero measure; $\ell > 0$ is the width of the common uniform distribution of X_j . Equivalently, one may write $\nu_N(s; \mathbf{J}^{(\ell)}; \omega)$ instead of $\nu_N(s; \mathbf{J}^{(\ell)}; X)$, since the sample space \mathbb{R}^N is identified with the underlying probability space Ω .

Since $\{X_i\}$ are IID with uniform distribution on $[0, \ell]$, the distribution of the random vector $X(\omega)$ is uniform in the cube $\mathbf{J}^{(\ell)} = [0, \ell]^N$, inducing a uniform conditional distribution on each element $\mathcal{X}(Y)$. Therefore, by (4.5) and (4.6),

$$\nu_N(s; \mathbf{J}^{(\ell)}) = \frac{N^{1/2}s}{|\mathcal{X}(Y)|}. \quad (4.7)$$

It is to be stressed that both sides of the above equality are random variables: $\nu_N(s; \ell) = \nu_N(s; \ell; \omega)$ by its definition in (4.6), and $\mathcal{X}(Y) = \mathcal{X}(Y(X(\omega)))$.

5. SHORT INTERVALS ARE UNLIKELY

Lemma 1. *Assume that the IID random variables X_1, \dots, X_N , $N \geq 2$, admit (common) probability density p_V with $\|p_V\|_\infty \leq \bar{\rho} < \infty$. Then*

$$\mathbb{P} \{ |\mathcal{X}(Y)| < r \} \leq \frac{1}{4} \bar{\rho}^2 r^2 N. \quad (5.1)$$

In particular, for $X_j \sim \text{Unif}([0, \ell])$, one has

$$\mathbb{P} \{ |\mathcal{X}(Y)| < r \} \leq \frac{r^2 N}{4\ell^2}. \quad (5.2)$$

Proof. Let

$$\underline{X} = \underline{X}(X) = \min_i X_i, \quad \overline{X} = \overline{X}(X) = \max_i X_i. \quad (5.3)$$

While $\overline{X}(X)$ and $\underline{X}(X)$ vary along the elements $\mathcal{X}(Y)$, their difference $\overline{X}(X) - \underline{X}(X)$ does not; it is uniquely determined by $\mathcal{X}(Y)$.

According to Remark 4.1, each $N^{1/2}X_i$, $i = 1, \dots, N$, restricted to $\mathcal{X}(Y)$, provides a normalized length parameter on $\mathcal{X}(Y)$; thus the range of each $N^{1/2}X_i|_{\mathcal{X}(Y)}$ is an interval of length $|\mathcal{X}(Y)|$. One can increase (resp., decrease), e.g., the value of X_1 , as long as *all* $\{X_i, 1 \leq i \leq N\}$ are strictly smaller than ℓ (resp., strictly positive). Therefore, the maximum increment of X_1 (indeed, of any X_i) along $\mathcal{X}(Y)$ is given by $\ell - \overline{X}(X)$, and its maximum decrement equals $\underline{X}(X)$, so the range of the normalized length parameter $N^{1/2}X_1$ along $\mathcal{X}(Y(X))$ is an interval of length $N^{1/2}(\ell - \overline{X}(X) + \underline{X}(X))$:

$$|\mathcal{X}(Y(X))| = N^{1/2}(\ell - \overline{X}(X) + \underline{X}(X)), \quad (5.4)$$

Since both $\underline{X}(X)$ and $\ell - \overline{X}(X)$ are non-negative,

$$\underline{X} + (\ell - \overline{X}) < t \implies \max\{\underline{X}, \ell - \overline{X}\} < t. \quad (5.5)$$

With $0 \leq t \leq \ell$, $(\ell - X_i < t/2)$ implies $(X_i > t/2)$, thus denoting

$$A_{ij}(t) := \{X_i < t/2\} \cap \{\ell - X_j < t\}, \quad (5.6)$$

we have, for any i ,

$$A_{ii}(t) = \{X_i < t\} \cap \{\ell - X_i < t\} = \emptyset. \quad (5.7)$$

Therefore,

$$\{\max\{\underline{X}(X), \ell - \overline{X}(X)\} < t\} \subset \bigcup_{i \neq j} \left\{X_i < \frac{t}{2}, \ell - X_j < t\right\}. \quad (5.8)$$

Thus the union $\cup_{i \neq j} A_{ij}(t)$ contains all samples X with $|\mathcal{X}(Y)| < t$.

The sample $\{X_k\}$ is IID, with common probability density uniformly bounded by $\bar{\rho} < \infty$, so for any $i \neq j$

$$\mathbb{P}\{A_{ij}(t)\} = \mathbb{P}\{X_i < t\} \cdot \mathbb{P}\{\ell - X_j < t\} = \bar{\rho}^2 t^2.$$

Therefore,

$$\begin{aligned} \mathbb{P}\{|\mathcal{X}(Y)| < r\} &= \mathbb{P}\left\{N^{1/2}((\ell - \overline{X}(X)) + \underline{X}(X)) < r\right\} \\ &= \mathbb{P}\left\{((\ell - \overline{X}(X)) + \underline{X}(X)) < rN^{-1/2}\right\} \\ &\leq \sum_{i \neq j} \mathbb{P}\left\{A_{ij}(rN^{-1/2})\right\} \leq N(N-1) (\bar{\rho} r N^{-1/2})^2 \\ &\leq \bar{\rho}^2 r^2 N. \end{aligned} \quad (5.9)$$

□

6. REGULARITY BOUND FOR THE UNIFORM DISTRIBUTIONS

Theorem 1. *Let be given IID random variables X_1, \dots, X_N with $X_i \sim \text{Unif}([0, \ell])$ and a measurable function $\lambda : Y \mapsto \lambda(Y)$. In each interval $\mathcal{X}(Y) \subset \tilde{\mathcal{X}}(Y) \cong \mathbb{R}$, introduce the sub-interval $I_s(Y) = [\lambda(Y), \lambda(Y) + s] \cap \tilde{\mathcal{X}}(Y)$. For any $s \in (0, 1]$,*

$$\mathbb{P}\{\xi(\omega) \in I_s(Y)\} \leq \frac{3N^3}{\ell} s. \quad (6.1)$$

Proof. Let $\mathfrak{l}(\omega) := |\mathcal{X}(Y)|$. The function ξ cannot serve as a normalized length parameter on the intervals parallel to $(1, \dots, 1)$, since its gradient $(1/N, \dots, 1/N)$ has norm $1/\sqrt{N}$. For this reason, it is convenient to introduce its normalized counterpart $\tilde{\xi} = \xi\sqrt{N}$ and rescaled intervals $\tilde{I}_s = [\tilde{\lambda}, \tilde{\lambda} + s\sqrt{N}]$, $\tilde{\lambda} = \lambda\sqrt{N}$.

$$\begin{aligned} \mathbb{P}\{\xi \in I_s(\eta)\} &= \mathbb{P}\{\tilde{\xi} \in \tilde{I}_s(\eta)\} = \mathbb{E}\left[\mathbb{P}\{\tilde{\xi} \in \tilde{I}_s(\eta) \mid \mathfrak{F}_\eta\}\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\mathfrak{l}(\omega) < s\sqrt{N}} \mathbb{P}\{\tilde{\xi} \in \tilde{I}_s(\eta) \mid \mathfrak{F}_\eta\}\right] + \mathbb{E}\left[\mathbf{1}_{\mathfrak{l}(\omega) \geq s\sqrt{N}} \mathbb{P}\{\tilde{\xi} \in \tilde{I}_s(\eta) \mid \mathfrak{F}_\eta\}\right] \\ &\leq \mathbb{P}\{\mathfrak{l}(\omega) < s\sqrt{N}\} + \mathbb{E}\left[\mathbf{1}_{\mathfrak{l}(\omega) \geq s\sqrt{N}} \mathbb{P}\{\tilde{\xi} \in \tilde{I}_s(\eta) \mid \mathfrak{F}_\eta\}\right] \end{aligned} \quad (6.2)$$

where, by virtue of (5.9),

$$\mathbb{P}\{\mathfrak{l}(\omega) < s\sqrt{N}\} \leq \frac{N^2}{\ell^2} s^2, \quad (6.3)$$

yielding

$$\sup_{s>0} \frac{\mathbb{P}\{\mathfrak{l}(\omega) < s\}}{s^2} \leq \frac{N^2}{\ell^2}. \quad (6.4)$$

The second summand in the RHS of (6.2) can be assessed as follows:

$$\begin{aligned} \mathbb{E} \left[\mathbf{1}_{\mathfrak{l} \geq s\sqrt{N}} \mathbb{P} \left\{ \tilde{\xi} \in \tilde{I}_s(\eta) \mid \mathfrak{F}_\eta \right\} \right] &\leq \mathbb{E} \left[\mathbf{1}_{\mathfrak{l} \geq s\sqrt{N}} \frac{s\sqrt{N}}{\mathfrak{l}} \right] = s\sqrt{N} \mathbb{E} \left[\mathbf{1}_{\mathfrak{l} \geq s} \mathfrak{l}^{-1} \right] \\ &= s\sqrt{N} \int_{s\sqrt{N}}^{\ell\sqrt{N}} r^{-1} dF_{\mathfrak{l}}(r) \end{aligned} \quad (6.5)$$

Using integration by parts for the Stieltjes integral and (6.4), we obtain

$$\begin{aligned} \int_{s\sqrt{N}}^{\ell\sqrt{N}} r^{-1} dF_{\mathfrak{l}}(r) &= \frac{F(r)}{r} \Big|_{s\sqrt{N}}^{\ell\sqrt{N}} + \int_{s\sqrt{N}}^{\ell\sqrt{N}} r^{-2} F_{\mathfrak{l}}(r) dr \\ &\leq \frac{1}{\ell\sqrt{N}} + \ell\sqrt{N} \sup_{r>0} \frac{F_{\mathfrak{l}}(r)}{r^2} \leq \frac{1}{\ell\sqrt{N}} + \frac{\ell\sqrt{N} \cdot N^2}{\ell^2} \\ &\leq \frac{2N^{5/2}}{\ell}. \end{aligned} \quad (6.6)$$

Collecting (6.3), (6.5) and (6.6), and taking into account that $s/\ell \leq 1$, the assertion follows:

$$\mathbb{P}\{\xi \in I_s(\eta)\} \leq \frac{N^2}{\ell^2} s^2 + \frac{2N^{5/2}}{\ell} s \leq \frac{3N^3}{\ell} s. \quad (6.7)$$

□

7. SMOOTH POSITIVE DENSITIES

Now we consider a richer class of probability distributions. While the conditions which we will assume are certainly very restrictive, they are quite sufficient for applications to physically realistic Anderson models.

Theorem 2. *Assume that the common probability distribution of the IID random variables V_j , $j = 1, \dots, N$, with PDF F_V , satisfies the following conditions:*

(i) *the probability distribution is absolutely continuous:*

$$dF_V(v) = \rho(v) dv, \quad \text{supp } \rho = [a, a + \ell]; \quad (7.1)$$

(ii) *the probability density $\rho(\cdot)$ has bounded logarithmic derivative on $(a, a + \ell)$:*

$$\|(\ln \rho)' \mathbf{1}_{(a, a+\ell)}\|_\infty \leq C'_\rho < +\infty. \quad (7.2)$$

Then there exists a constant $C = C(F_V, \ell) < \infty$ such that for any $s \in (0, \ell N^{-2})$ and any \mathfrak{F}_η -measurable random variable λ , setting $I_s(\omega) := [\lambda(\omega), \lambda(\omega) + s]$, one has the following bound:

$$\mathbb{P}\{\xi_N(\omega) \in I_s(\omega)\} \leq CNs. \quad (7.3)$$

Proof. Without loss of generality, it suffices to prove the claim for $\text{supp } \rho = [0, \ell]$, which we assume below.

◆ As in Section 3, introduce a partition of the sample space into the cubes \mathbf{J}_k , induced by the decomposition $[0, \ell] = \sqcup_k J_k$,

$$J_k = \left[\frac{k-1}{M_N}, \frac{k}{M_N} \right], \quad k = 1, \dots, M_N = N^2.$$

We have then

$$\mathbf{J}_{\mathbf{k}} = \bigtimes_{i=1}^N J_{k_i}, \quad \mathbf{k} = (k_1, \dots, k_N).$$

◆ The hypothesis (7.2) implies that for any $\mathbf{x} \in \mathbf{J}_{\mathbf{k}}$ the logarithm of $\mathbf{p}(\mathbf{x})$ is well-defined and satisfies

$$|\ln \mathbf{p}(\mathbf{x}) - \ln \mathbf{p}(\mathbf{a}_{\mathbf{k}})| \leq \sum_{i=1}^N |\ln \rho(x_i) - \ln \rho(a_{k_i})| \leq N C'_p \ell M_N^{-1} = O(\ell N^{-1}).$$

thus, setting $\alpha_N = \ell N^{-1}$,

$$\forall \mathbf{x} \in \mathbf{J}_{\mathbf{k}} \quad \frac{\mathbf{p}(\mathbf{x})}{\mathbf{p}(\mathbf{a}_{\mathbf{k}})} \in [e^{-\alpha_N}, e^{+\alpha_N}].$$

Now introduce in $\mathbf{J}_{\mathbf{k}}$:

- the uniform probability distribution $\tilde{\mathbf{P}}_{\mathbf{k}}$, i.e., the normalized measure with constant density $\tilde{\mathbf{p}}_{\mathbf{k}}$ w.r.t. the Lebesgue measure;
- the probability distribution induced by \mathbf{P} , conditional on $\{\mathbf{X} \in \mathbf{J}_{\mathbf{k}}\}$, i.e., the normalized measure with density

$$\mathbf{p}_{\mathbf{k}}(\mathbf{x}) = Z_{\mathbf{k}}^{-1} \mathbf{p}(\mathbf{x}) = \frac{\mathbf{p}(\mathbf{x})}{\int_{\mathbf{J}_{\mathbf{k}}} \mathbf{P}(\mathbf{y}) d\mathbf{y}}$$

By continuity of the density \mathbf{p} , $\int_{\mathbf{J}_{\mathbf{k}}} \mathbf{P}(\mathbf{y}) d\mathbf{y} = c |\mathbf{J}_{\mathbf{k}}|$, for some $c \in [e^{-\alpha_N}, e^{+\alpha_N}]$, so

$$\frac{\mathbf{p}_{\mathbf{k}}(\mathbf{x})}{\tilde{\mathbf{p}}(\mathbf{x})} = \frac{\mathbf{p}(\mathbf{x})}{c} \in [e^{-2\alpha_N}, e^{+2\alpha_N}]$$

Hence for any event \mathcal{A} , we have

$$e^{-2\alpha_N} \mathbb{P}\{\mathcal{A}\} \leq \mathbb{P}_{\mathbf{k}}\{\mathcal{A}\} \leq e^{+2\alpha_N} \mathbb{P}\{\mathcal{A}\} \quad (7.4)$$

◆ It follows from (7.4) and (3.1) that

$$\mathbb{P}\{\xi \in I_s(\eta)\} \leq \sup_{\mathbf{k}} \mathbb{P}_{\mathbf{k}}\{\xi \in I_s(\eta)\} \leq C(F_V, \ell) N s. \quad (7.5)$$

Recall that this bound was proved only for $s \leq \ell/M(N) = o(\ell N^{-1})$.

□

REFERENCES

- [1] V. Chulaevsky, *A remark on charge transfer processes in multi-particle systems*, 2010, [arXiv:math-ph/1005.3387](#). ↑
- [2] V. Chulaevsky, *On resonances in disordered multi-particle systems*, C.R. Acad. Sci. Paris, Ser. I, **350** (2011), 81–85. ↑
- [3] F. Wegner, *Bounds on the density of states in disordered systems*, Z. Phys. B. Condensed Matter **44** (1981), 9–15. ↑

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE REIMS, MOULIN DE LA HOUSSE, B.P. 1039, 51687 REIMS CEDEX 2, FRANCE, E-MAIL: VICTOR.TCHOULAEVSKI@UNIV-REIMS.FR